Contribution ID: **142** Type: **Talk**

Lattice QCD on the NVIDIA Grace-Hopper and Blackwell Architectures

Tuesday 4 November 2025 17:40 (20 minutes)

The evolution of GPU architectures has fundamentally transformed computational capabilities for Lattice QCD simulations. Over recent generations, GPU memory bandwidth has increased by approximately one order of magnitude, with modern systems achieving terabytes per second of memory throughput, while computational width has expanded significantly. These advances, driven primarily by artificial intelligence workload demands, present new optimization opportunities for Lattice QCD calculations.
We present comprehensive optimization strategies implemented in QUDA that fully exploit the potential of modern single-GPU systems. Additionally, we examine the benefits of coherent memory architectures in Grace-Hopper and Grace-Blackwell superchip platforms. Finally, we investigate strong scaling properties of Lattice QCD applications on multi-node NVLink Grace-Blackwell systems.

Parallel Session (for talks only)

Software development and machines

Authors: Dr JOO, Balint (NVIDIA); WEINBERG, Evan (NVIDIA); Dr TU, Jiqun (NVIDIA); CLARK, Kate (NVIDIA); WAGNER, Mathias (Nvidia)

Presenter: WAGNER, Mathias (Nvidia)

Session Classification: Software development and machines